

# Enhanced Non-Gaussianity from Excited Initial States

---

**R. Holman**

*Department of Physics, Carnegie Mellon University  
Pittsburgh PA 15213 USA  
E-mail: rh4a@andrew.cmu.edu*

**Andrew J. Tolley**

*Perimeter Institute for Theoretical Physics  
31 Caroline St. N, Waterloo ON N2L 2Y5 Canada  
E-mail: atolley@perimeterinstitute.ca*

**ABSTRACT:** We use the techniques of effective field theory in an expanding universe to examine the effect of choosing an excited inflationary initial state built over the Bunch-Davies state on the CMB bi-spectrum. We find that even for Hadamard states, there are unexpected enhancements in the bi-spectrum for certain configurations in momentum space due to interactions of modes in the early stages of inflation. These enhancements can be parametrically larger than the standard ones and are potentially observable in future data. These initial state effects have a characteristic signature in  $l$ -space which distinguishes them from the usual contributions, with the enhancement being most pronounced for configurations corresponding to flattened triangles for which two momenta are collinear.

---

## Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Effective Field Theory in an Expanding Universe</b>	<b>6</b>
<b>3. The Inflationary Three Point Function</b>	<b>10</b>
3.1 Scalar Minimally Coupled to Gravity	12
3.2 Higher derivative interactions	14
3.3 Backreaction from higher derivative interactions	18
<b>4. Measure issues and the l-space bi-spectrum</b>	<b>21</b>
<b>5. Higher irrelevant operators, N-point functions</b>	<b>24</b>
<b>6. Conclusions</b>	<b>25</b>

---

## 1. Introduction

Inflation has become the dominant paradigm for the study of the early universe. Current data from the CMB power spectrum [1], as well as large scale structure [2] are certainly consistent with the assumption of an inflationary phase in the early universe, while the anti-correlation between the TT and TE power spectra on superhorizon scales should soon become strong enough statistically to serve as a “smoking gun” for inflation. Furthermore, the case for inflation will certainly be strengthened by the discovery of the B-mode polarization of the CMB.

These cosmological observations provide a powerful lever arm that allows us access to physics at energy scales far beyond those that will be probed by accelerators in the foreseeable future. Already, we can use measurements of the spectral index  $n_s$  and bounds on the ratio of tensor to scalar fluctuations  $r$  to place restrictions on the form of the inflaton scalar potential, at least within the field range corresponding to the observable 10 or so e-folds of inflation [1].

The inflaton potential is only part of the story, however. The inflaton should almost certainly be viewed as an effective degree of freedom, perhaps arising from a

higher dimensional theory as the separation mode between a brane and an antibrane [3], or as one of the many axions that appears in string theory [4], amongst many other possibilities. As such, the inflaton effective action will contain not only renormalizable terms, but also so-called irrelevant operators which are suppressed by powers of the cutoff scale  $M$ . At energies higher than  $M$  this effective theory breaks down and the inflaton should not be viewed as the appropriate degree of freedom. The effect of these operators needs to be taken into account when computing the various inflationary observables, and several works have considered these corrections [5]. Whilst these corrections might a priori have been expected to be small, they can give rise to corrections to the CMB bi-spectrum which are larger than the standard result as shown by Creminelli [6]. This arises because the standard contribution to the bi-spectrum is already highly suppressed by the slow roll parameters [7, 8].

Equally as important as including higher order terms in the effective action is a proper treatment of the quantum state of the fluctuations of the inflaton about its zero mode. The standard calculations of the power spectrum are predicated on a particular choice of quantum state, the so-called Bunch-Davies (BD) state [9]. There has been a great deal of work in recent years focussed on understanding what might constitute reasonable modifications to this quantum state and how the data we already have on inflationary observables from the power spectrum might constrain these modifications [10].

In this work we ask the question: can we probe the nature of the initial quantum state of the inflaton through the non-gaussianities produced via inflaton fluctuations? We will argue that statistics that probe this non-gaussianity are amazingly sensitive to the nature of both the interactions of the inflaton and more importantly, to its initial state. In fact, in some ways, higher correlation functions are more sensitive probes of initial state fluctuations than the power spectrum. The reason for this is that the power spectrum is only sensitive to the interactions of the inflaton through loops which are highly suppressed [11, 13] (see also [12] for slightly different conclusions). On the other hand, non-gaussianities can probe the inflaton's interactions directly at tree-level, which can be significant if the inflaton does not start out in the BD vacuum state.

The simplest correlation function that probes the non-gaussian nature of inflaton statistics is the three point correlation function, otherwise known as the bi-spectrum, of the fluctuations. The seminal calculation of the bi-spectrum was done by Maldacena [7] (see also [8]) who expanded the action for a minimally coupled inflaton scalar coupled to gravity to third order in the gauge invariant curvature fluctuation variable  $\zeta$ . This is the relevant quantity since it remains constant outside the horizon, re-emerging as a non-gaussian contribution to perturbations on scales relevant to present day cosmology. Subsequent calculations have refined estimates of the bi-spectrum in the standard case,

extended them to other inflationary models [14], and have also considered the effect of irrelevant operators [6].

Let us briefly sketch the details of the calculations. In a cosmological setting, we need to consider the time evolution of equal time correlation functions. Thus we choose an initial state at a (conformal) time  $\eta_0$ , which we take to be at or near the onset of the inflationary phase. There is considerable freedom in how to choose this state and we will discuss this issue further below. To compute the time evolution, we work in the interaction picture where the state evolves according to

$$i \frac{d|\psi(\eta)\rangle}{d\eta} = H_I |\psi(\eta)\rangle, \quad (1.1)$$

where  $H_I$  is the interaction Hamiltonian. This has the formal solution given an initial state  $|\psi(\eta_0)\rangle$  at a (finite) time  $\eta = \eta_0$ :

$$|\psi(\eta)\rangle = T e^{-i \int_{\eta_0}^{\eta} H_I(\eta') d\eta'} |\psi(\eta_0)\rangle. \quad (1.2)$$

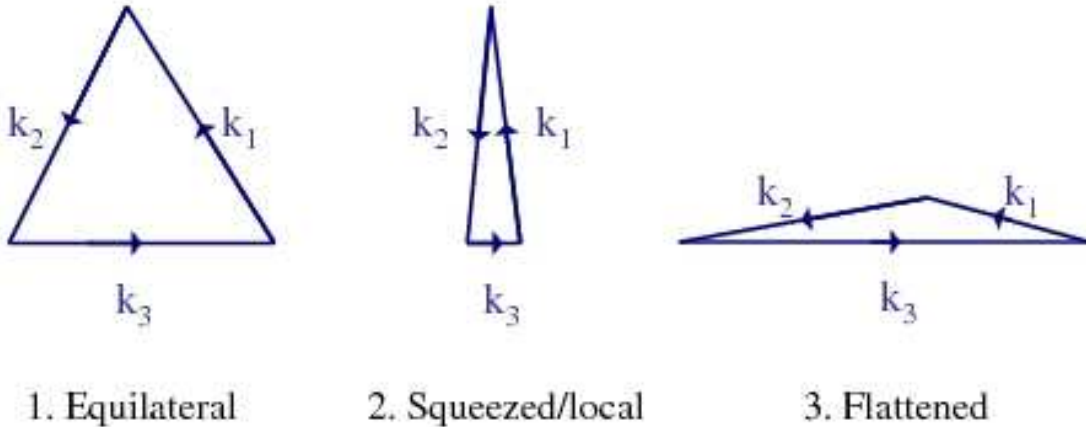
We require that our effective theory be valid at  $\eta = \eta_0$ . In particular, as will be discussed in Sec. 2, this means that states are not excited at momentum scales greater than the cutoff. The initial state can be described by giving all of the correlation functions of our dynamical degrees of freedom, e.g. the comoving curvature perturbation  $\zeta(\eta)$ . In particular the tree level contribution to the equal time three point function is given by

$$\begin{aligned} \langle \psi(\eta) | \zeta_{\vec{k}_1}(\eta) \zeta_{\vec{k}_2}(\eta) \zeta_{\vec{k}_3}(\eta) | \psi(\eta) \rangle &= \langle \psi(\eta_0) | \zeta_{\vec{k}_1}(\eta) \zeta_{\vec{k}_2}(\eta) \zeta_{\vec{k}_3}(\eta) | \psi(\eta_0) \rangle \\ &- i \int_{\eta_0}^{\eta} d\eta' \langle \psi(\eta_0) | [\zeta_{\vec{k}_1}(\eta) \zeta_{\vec{k}_2}(\eta) \zeta_{\vec{k}_3}(\eta), H_I(\eta')] | \psi(\eta_0) \rangle + \mathcal{O}(H_I^2). \end{aligned} \quad (1.3)$$

Thus the total three point correlator is a sum of a contribution from any initial non-gaussianity present in the state at the beginning of inflation, evolved forward in time with the free Hamiltonian, and the contribution that occurs from interactions that take place between the beginning of inflation and the time of observation. The former contribution has to be treated as arbitrary until we have a better understanding of the pre-inflationary stage. At best, we can place bounds on its contribution based on ensuring that its backreaction is small. The second contribution, on the other hand, depends explicitly on the interactions of the theory as well as the initial state. This contribution will be the focus of our investigations in this work.

In practice we shall assume that the initial state is gaussian, so that all the correlation functions of operators in this state will be completely specified by the two point Wightman function (non time-ordered in-in expectation value), which in momentum space takes the form

$$\langle \psi(\eta_0) | \zeta_{\vec{k}_1}(\eta) \zeta_{\vec{k}_2}(\eta') | \psi(\eta_0) \rangle = (2\pi)^3 \delta^3(\vec{k}_1 + \vec{k}_2) G_{\vec{k}_1}^>(\eta, \eta') = (2\pi)^3 \delta^3(\vec{k}_1 + \vec{k}_2) \mathcal{V}_{\vec{k}_1}(\eta) \mathcal{V}_{\vec{k}_1}^*(\eta'), \quad (1.4)$$



**Figure 1:** Momentum triangles which dominate the bi-spectrum: 1. Equilateral triangles arising from Hubble-horizon crossing. 2. Squeezed or local triangles arising from super-Hubble evolution. 3. Flattened triangles arising from initial state effects.

where  $\mathcal{V}_k$  are properly normalized solutions of the linear equations of motion for  $\zeta$  in momentum space. In Sec. 2 we discuss the conditions required of this two-point function so that we can reliably trust our calculation.

In the standard calculation using the BD vacuum, the dominant contribution to the three point function, which measures directly three particle interactions, comes from when the modes cross the horizon. The intuitive reason for this is that at sub-horizon scales the BD vacuum corresponds to a state of no-particles<sup>1</sup> (where the particles are inflaton quanta). Thus there are no particles to interact and so no contribution to the three point function. As the modes cross the horizon, the WKB approximation breaks down which is tantamount to the statement that particles are created. These particles can then undergo interactions which contribute to the three point function. Once the modes are well outside the horizon,  $\zeta$ , properly defined, gets frozen in nonlinearly i.e. it is conserved and further interactions become irrelevant. In curvaton models [15] there can be an additional contribution from super-horizon scales because  $\zeta$  is not necessarily conserved [16]. In figure 1 we show the momentum triangles for which the three point function is maximized for these two types of effects (triangles 1 and 2).

The main point of this paper is that this situation changes dramatically if the initial state is not BD. In this case there are particles present initially which can undergo

---

<sup>1</sup>It is true that for a static observer the BD vacuum appears as a thermal bath of particles. However here we are using the adiabatic or WKB definition of particles appropriate to the flat slicing of de Sitter. This is the definition that is actually most useful in the context of calculations.

interactions. Furthermore the interactions of the inflaton are necessarily stronger at the beginning of inflation than at the end, and so we get a second contribution to the non-gaussianity coming from interactions in the early stages of inflation. These effects dominate the bi-spectrum for the flattened triangles (defined as those for which two momentum are collinear) of type 3 in figure 1 (these were called ‘folded’ in Ref. [17] where a similar study was performed). To see why interactions get stronger in the past, let us consider the prototypical example of an interacting scalar field  $\phi$  on an FRW geometry  $ds^2 = a^2(\eta)(-d\eta^2 + d\vec{x}^2)$  with Lagrangian expressed in conformal time

$$S = \int d\eta \int d^3x a^4 \left( -\frac{1}{2a^2}(\partial\phi)^2 - \sum_{n=2}^{\infty} \frac{\lambda_n}{n!M^{n-4}}\phi^n - \sum_{m=2}^{\infty} \frac{k_m}{M^{4m-4}} \left( \frac{1}{a^2}(\partial\phi)^2 \right)^m + \dots \right), \quad (1.5)$$

where  $M$  is the cutoff scale and the ellipsis corresponds to additional higher derivative interactions expected on usual grounds. Upon rescaling to the canonically normalized ‘comoving’ field  $\chi = a\phi$  we obtain<sup>2</sup>

$$S = \int d\eta \int d^3x \left( -\frac{1}{2}(\partial\chi)^2 + \frac{a''}{2a}\chi^2 - \sum_{n=2}^{\infty} \frac{\lambda_n}{n!a^{n-4}M^{n-4}}\chi^n - \sum_{m=2}^{\infty} \frac{k_m}{M^{4m-4}} \frac{1}{a^{4m-4}} \left( (\partial\chi)^2 + \frac{a'^2}{a^2}\chi^2 - 2\frac{a'}{a}\chi\chi' \right)^m + \dots \right). \quad (1.6)$$

Note that the scaling with  $a$  is the same as that with the cutoff  $M$ . At subhorizon scales the terms suppressed by  $a'/a$  and  $a''/a$  are irrelevant. It is apparent that higher order potential interactions with  $n > 4$  and higher order kinetic interactions with  $m \geq 2$  necessarily grow in the past as  $a \rightarrow 0$ . In other words the irrelevant operators become more relevant as modes become blue-shifted. In this case these terms will not contribute to the three point function at tree level (the first interesting contribution is to the 4-pt from the kinetic interactions) but as we shall see later when expanding around the background inflaton solution  $\phi_0(\eta)$ , we easily generate contributions to the three point function that scale as positive powers of  $1/a$ . Somewhat to our surprise then, it appears that the theory with higher derivative terms gives rise to a greater enhancement than the theory without them; the standard expectation is that irrelevant operators would be subdominant for low energy observables such as the CMB bi-spectrum<sup>3</sup>.

The upshot of our calculation is that the bi-spectrum is *at least* as good a probe of the initial state of inflaton fluctuations as the power spectrum. The effects in both

---

<sup>2</sup>In the full gauge invariant calculation this is the familiar Mukhanov variable  $v$ .

<sup>3</sup>This is a reflection of the fact that higher derivative operators are irrelevant *in the IR*. However, the scales that are in the IR today, relative to the cutoff scale  $M$ , were in the UV, or at least less in the IR at earlier times due to the expansion of the Universe.

cases will depend on the Bogoliubov  $\beta_k$  coefficient (see below) but there is *no* extra enhancement for the case of the power spectrum. We also see that a great deal of physics can be missed by not including the effect of irrelevant operators. Earlier works [18] have also considered the effects of initial state effects but not directly their implications for interactions. More recently in Ref. [19], similar but less explicit conclusions were made about the non-gaussianities. However, our approach to considering the typical corrections and in particular the magnitude of the backreaction [20] is somewhat different. The general features of the effects considered here including the precise shape dependence were pointed out in Ref. [17], although explicit estimates of the possible magnitude of these effects were not given.

In Sec. 2 we discuss how effective field theories are to be construed in the context of an expanding universe, and give a preliminary discussion of the importance of bounding backreaction. We then turn to our calculations of the bi-spectrum for the renormalizable and higher derivative interactions of the form described above in Sec. 3. We will see that for certain triangles in momentum space, there are additional enhancements as a direct consequence of the inflaton interactions taking place at the beginning of inflation. In Sec. 4 we then show that these effects have sufficient measure, so that the enhancement survives even after the bi-spectrum is converted to spherical harmonic space. This is where the higher derivative operators triumph over the renormalizable ones, even when the effects in both theories are calculated with the same excited states. In Sec. 5 we discuss the form of the enhancements from higher irrelevant operators and finally we conclude in Sec. 6.

## 2. Effective Field Theory in an Expanding Universe

The usual approach to effective field theories [21] consists of the following steps. First identify applicable regime of energies/temperatures in which we want to do physics. Next, decide what the relevant degrees of freedom in this regime should be, as well as what symmetries their dynamics should obey. Finally write down the most general lagrangian in terms of these degrees of freedom that incorporates the required symmetries. Once this is done, observables of the theory can be calculated in this regime. In this calculation, a determination of how accurate the results must be (to match experimental data, say) tells us which terms and in particular, which irrelevant operators to keep in our action.

While this approach is sufficient for standard particle physics observables, it is deficient when it comes to calculating in an expanding universe. In this situation, we are not calculating S-matrix elements (which may not even exist in some cosmologies, such as de Sitter [22]) which depend on boundary values at asymptotic times. Instead, we

need to compute the time evolution of correlation functions which requires us to compute so-called “in-in” matrix elements [23], and which incorporates the time evolution of the states, as well as the operators. In essence, we are solving a initial-value problem rather than a boundary one.

The notion of an effective theory presupposes an energy scale  $M$  such that we are constrained to do physics only below this scale. The new twist arising from the universes’ expansion is the fact that it induces a redshift of energies so that scales that were once larger than  $M$  will eventually become part of the low energy ( $< M$ ) spectrum. In fact, we can turn this around to make the following statement. Consider a *physical* momentum scale  $k$  corresponding to a length scale on the CMB sky today. Then, at a (conformal) time  $\eta_0$  such that  $k/a(\eta_0) \sim M$  we reach the limit of validity of the effective theory for this scale. If we want our effective theory to be valid for all of the relevant scales in the CMB sky, we must then impose a limit on how far back in time we can trust this theory; we will let  $\eta_0$  denote the earliest time at which the effective field theory can be trusted, and we shall take this time to be the beginning of inflation, although more generally we only require it to be the time at which the physical scales of the CMB today are of the order of the cutoff.

The scale  $M$  can also infiltrate the description of the initial state of the fluctuations. The standard way this state is chosen is by arguing that at short enough distances, the fluctuations behave as if they are in flat space, and their vacuum state would reflect this. In practice, the  $\eta \rightarrow -\infty$  limit of the solutions to the mode equation is taken, and the linear combination of the solutions that approaches a positive energy plane wave in this limit is then chosen. The state picked by this procedure in the case of a de Sitter universe is the Bunch-Davies (BD) state [9]; we will denote these modes via  $\mathcal{U}_k$  in the following.

Given that at times earlier than  $\eta_0$  it may not be permissible to treat the inflaton as a well defined degree of freedom and even if we could, we would certainly *not* be privy to the relevant dynamics, it does not seem reasonable to use a state whose definition required going to arbitrarily short distances. In fact, using this state requires the radical assumption that the description of the physics in terms of the inflaton fluctuations as a free scalar field in an FRW background is a valid one at all scales.

A more reasonable description of the state should depend more explicitly on the domain of validity of the theory. Thus we set initial conditions at  $\eta_0$  and write

$$\mathcal{V}_k(\eta) = \alpha_k \mathcal{U}_k(\eta) + \beta_k \mathcal{U}_k^*(\eta), \quad |\alpha_k|^2 - |\beta_k|^2 = 1, \quad (2.1)$$

*i.e.* the mode  $\mathcal{V}_k$  is a Bogoliubov transform of the BD state. As is well known, this can be viewed as an excited state built upon the BD state; the number density of particles



of momentum  $k$  is  $|\beta_k|^2$ . The Bogoliubov coefficients encode information about the initial conditions satisfied by the  $\mathcal{V}_k$  relative to the BD modes: if we specify that

$$\dot{\mathcal{V}}_k(\eta_0) = -i\varpi_k \mathcal{V}_k(\eta_0), \quad \dot{\mathcal{U}}_k(\eta_0) = -i\omega_k \mathcal{U}_k(\eta_0), \quad (2.2)$$

then we see that

$$\frac{\beta_k}{\alpha_k} = \frac{\omega_k - \varpi_k}{\omega_k + \varpi_k}. \quad (2.3)$$

How should we fix  $\beta_k$  (or equivalently,  $\varpi_k$ )? In Ref. [24]  $\varpi_k$  is described in terms of a Laurent expansion in  $\omega_k/M$ , and a new renormalization procedure has to be implemented to account for the spacelike boundary on which the initial conditions are specified. In this work we follow the more standard approach and demand that the state constructed from the modes  $\mathcal{V}_k$  be Hadamard [25]. This fixes the short-distance behavior to be the usual one, namely that  $\beta_k$  must fall off faster than  $1/k^2$ , but it may otherwise be arbitrary. We implement this by noticing that since our description in terms of an effective theory breaks down at the scale  $M$ , we should not excite any modes with energies higher than this. Thus we demand that  $\beta_k \rightarrow 0$  for  $k > Ma(\eta_0)$ . Note that this assumption shields us from the transplankian problem [26] since as the universe expands and transplankian modes redshift to cisplankian scales, they enter in their vacuum state  $\beta_k = 0$ , and so only the subsequent cisplankian dynamics will be important.

## Backreaction

We can put some bounds on how large the non-vanishing  $\beta_k$ 's can be by considering the issue of backreaction. This was treated within the context of the standard renormalization procedure in Ref. [27] and in Ref. [20, 24, 28] within the context of boundary renormalization. The basic issue is whether the energy density coming from the “particles” of the BD vacuum contained in the initial state might overwhelm that of the inflaton zero mode, and thus prevent inflation from occurring.

This energy density can be estimated by taking the crude model  $\beta_k \sim \beta_0 e^{-k^2/(Ma(\eta_0))^2}$  and demanding that the energy density of the nearly massless quanta of the inflaton be less than  $M_{\text{pl}}^2 H^2$ :

$$\rho \sim \frac{1}{a^4} \int \frac{d^3 k}{(2\pi)^3} |\beta_k|^2 k \sim \frac{a(\eta_0)^4}{a(\eta)^4} |\beta_0|^2 M^4 \lesssim M_{\text{pl}}^2 H^2 \Rightarrow |\beta_0|^2 \lesssim \frac{M_{\text{pl}}^2 H^2}{M^4}. \quad (2.4)$$

More generally we can imagine any smooth fall-off as considered in Ref. [27], where a fall-off of the form  $|\beta_k|^2 \sim \mathcal{O}(k^{4+\delta})$  is assumed, and  $\delta > 0$  to ensure the Hadamard condition. This gives rise to similar results (up to factors of  $\mathcal{O}(1)$ ), and so for the purposes of this article we shall consider the above crude model.

The stress-energy tensor also contains contributions from interactions, and these contributions are non-zero as soon as we generate some non-gaussianities. We shall perform more explicit calculations of these contributions in Sec. 3.3, but it is straightforward to estimate the bounds on these effects. On dimensional grounds, as long as we remain in the regime of effective field theory, the largest contribution to the energy density (expanded in powers of  $\beta_0$ ) can be of order  $|\beta_0|^n M^4$  for  $n \geq 1$ . For  $|\beta_0| \leq 1$  we naively expect the terms linear in  $\beta_0$  to dominate. However, any term which is odd in  $\beta_0$  will also contain some powers of  $e^{2ik\eta}$  inside the momentum integrals. The same situation already arises in free theory where we get a contribution to the energy density of the form

$$\Delta\rho = -\frac{1}{a^4} \int \frac{d^3k}{(2\pi)^3} k \beta_k^* e^{2ik\eta} + c.c.. \quad (2.5)$$

The crucial point is that at early times, i.e. large  $\eta$ , the rapid oscillations of the exponential damp its contribution. For instance for the model  $\beta_k \sim \beta_0 e^{-k^2/(Ma(\eta_0))^2}$ , this contribution behaves as  $e^{-4\eta^2 M^2 a(\eta_0)^2} = e^{-4M^2 a(\eta_0)^2/(Ha(\eta))^2}$ . At early times the exponential is negligible, and once the spacetime has inflated to the point that  $a(\eta)/a(\eta_0) \approx M/H$  and the exponential is of  $\mathcal{O}(1)$ , the  $1/a^4$  scaling suppresses the energy density to be of order  $H^4$  and consequently negligible. Thus even with interactions included, the backreaction to the energy density can be no larger than  $|\beta_0|^2 M^4$ . This is borne out by the more detailed calculation in Sec. 3.3.

We must also make sure that the slow roll conditions are not violated. Since  $\dot{H} = -\epsilon H^2 = -\frac{1}{2M_{\text{mpl}}^2} (p + \rho)$  and  $\ddot{H} = 2\epsilon\eta' H^3 = -\frac{1}{2M_{\text{mpl}}^2} (\dot{p} - 3H(p + \rho))$ , where  $\epsilon$  and  $\eta'$  are implicitly defined slow roll parameters, then assuming  $\Delta\rho \sim |\beta_0|^2 M^4$ ,  $\Delta p \sim |\beta_0|^2 M^4$  and  $\Delta\dot{p} \sim |\beta_0|^2 H M^4$  we have the bounds

$$|\beta_0| \leq \sqrt{\epsilon} \frac{H M_{\text{pl}}}{M^2}, \quad (2.6)$$

$$|\beta_0| \leq \sqrt{\epsilon\eta'} \frac{H M_{\text{pl}}}{M^2}. \quad (2.7)$$

This conclusion, also reached in [28], differs strongly from the conclusions of [19, 20]. That author gives more pessimistic estimates based on the order  $\beta_k$  contributions which neglect the oscillating nature of the exponentials. It is also assumed that  $\Delta\dot{p} \sim M\Delta p$  which again is not the case because of the oscillatory damping; rather we obtain  $\Delta\dot{p} \sim H\Delta p$ .

We see that backreaction does not necessarily force us to a small value for  $|\beta_0|$ . If  $M \sim M_{\text{pl}}$ , then  $|\beta_0| \sim H/M_{\text{pl}} \sim 10^{-6}$ , while if the new physics is at scales *smaller* than  $M_{\text{pl}}$ , we can get a large value for  $|\beta_0|$ . In general we require  $H < M$  for inflation to be described within the regime of effective field theory, and we must also require

that  $M \leq M_{\text{pl}}$ . Nevertheless for the reasonable choice  $M \sim 10^{-4} M_{\text{pl}}$  we can obtain  $|\beta_0| \sim 1$  for  $\epsilon, \eta' \sim 10^{-2}$ . There are bounds on how large  $\beta_k$  can be coming from direct observations of the power spectrum, but for sufficiently small or weakly  $k$ -dependent  $\beta_k$  we can easily evade these [29].

To summarize: we treat inflaton fluctuations as being described by an effective field theory valid at momenta and energies below a scale  $M$ . This forces us to the notion of an earliest time  $\eta_0$  at which we can trust this theory. We then argue that a more natural set of states to use to describe inflaton fluctuations are Bogoliubov transforms of the BD state and we ensure that they remain Hadamard by cutting off the Bogoliubov coefficient  $\beta_k$  for  $k > Ma(\eta_0)$ . These coefficients are further constrained by the requirement that the backreaction on the inflating background is negligible. Our assumption that the states are Hadamard guarantees renormalizability (in the effective field theory sense). We now turn to the calculation of the bi-spectrum using these states.

### 3. The Inflationary Three Point Function

Our goal in this section is to show what impact the modifications to the quantum state of inflaton fluctuations described in Sec. 2 have on the momentum space (as well as  $\ell$ -space) structure of the three point function. To study fluctuations about the FRW background, we follow Maldacena [7] and use the ADM foliation of the spacetime as

$$ds^2 = -N^2 dt^2 + h_{ij} (dx^i + N^i dt) (dx^j + N^j dt), \quad (3.1)$$

where  $N$  is the lapse function and  $N^i$  the shift function. The gauge invariance of the Einstein action needs to be fixed and this can be done in one of several ways. For instance we can set the fluctuations in the inflaton field to zero  $\delta\Phi = 0$  and parametrize the metric fluctuations as:

$$h_{ij} = \exp(2Ht + 2\zeta) \hat{h}_{ij}, \quad \det \hat{h} = 1. \quad (3.2)$$

The curvature fluctuation  $\zeta$  is useful since it is a gauge invariant quantity and remains constant outside the horizon. Note that this definition of  $\zeta$  differs by a sign from that often found in the literature and this is relevant to the sign of the three point function since the bounds on  $f_{NL}$  are not symmetric [30]. For superhorizon wavelengths, gradients in  $\zeta$  can be neglected so that the  $\zeta$  can be viewed as shift in the time variable. Maldacena's calculation involved using this gauge and parametrization then expanding the Einstein action to cubic order in  $\zeta$ .

Another gauge that can be quite useful moves all the scalar fluctuations into the inflaton field:

$$\Phi(\vec{x}, t) = \phi(t) + \delta\phi(\vec{x}, t), \quad h_{ij} = \exp(2Ht) \hat{h}_{ij}. \quad (3.3)$$

In this gauge, we compute the bi-spectrum by first calculating the three point function of the fluctuations  $\delta\phi$ ,  $\langle\delta\phi\delta\phi\delta\phi\rangle$  and then relating this to  $\langle\zeta\zeta\zeta\rangle$ ; it is this latter quantity that encodes the observable non-gaussianities in the CMB. To go from  $\delta\phi$  to  $\zeta$  we have to take into account the non-linear evolution of  $\delta\phi$  outside the horizon which involves going to second order in the perturbations. This gives rise to a non-linear term of the form  $\langle\delta\phi\delta\phi\rangle\langle\delta\phi\delta\phi\rangle$  to  $\langle\zeta\zeta\zeta\rangle$  which serves to cancel the time dependence outside the horizon coming from the inflaton fluctuations so that  $\zeta$  remains time independent. To lowest order in slow roll parameters, which suffices for our needs, the relation between  $\delta\phi$  and  $\zeta$  is

$$\zeta = -\frac{H}{\dot{\phi}}\delta\phi + \mathcal{O}(\epsilon, \eta) \left( \frac{H}{\dot{\phi}}\delta\phi \right)^2, \quad (3.4)$$

where  $\epsilon, \eta$  are the slow roll parameters. We will use this latter gauge for our calculations, and only compute the leading order behavior in both the slow roll parameters, as well as in  $H/M$ . This means in particular that the fluctuation modes will be those defined on de Sitter.

As discussed in Sec. 2, in a cosmological setting, the relevant expectation values are of the in-in type, as opposed to the in-out ones used to compute S-matrix elements. There is a formalism in place for doing this [23] and it has been further elaborated in Refs. [11, 24, 31, 32]. We use these techniques here to compute the three point function.

In the canonical version of this formalism used by Maldacena, we first construct the interacting Hamiltonian,  $H_I$ , in the usual manner. The equal time tree level contribution to the three point function is then given by

$$\langle\zeta(x_1)\zeta(x_2)\zeta(x_3)\rangle = -i \int_{\eta_0}^{\eta} d\eta' \langle[\zeta(x_1)\zeta(x_2)\zeta(x_3), H_I(\eta')]\rangle_0; \quad (3.5)$$

using the reality condition of the equal time product this is the same as

$$\langle\zeta(x_1)\zeta(x_2)\zeta(x_3)\rangle = -2 \mathcal{R}e \left( \int_{\eta_0}^{\eta} d\eta' i \langle\zeta(x_1)\zeta(x_2)\zeta(x_3)H_I(\eta')\rangle_0 \right). \quad (3.6)$$

The lower limit on the integrals corresponds to the beginning of inflation  $\eta_0$ . This is not an artificial cutoff, but rather it reflects our choice of a *gaussian* initial state. In practice, this cutoff is crucial if we are to make sense of the integrals when initial states other than the Bunch-Davies one are chosen.

The free field correlators  $\langle \zeta(x_1)\zeta(x_2)\zeta(x_3)H_I(\eta') \rangle_0$  can be computed via Wick's theorem, where contractions are replaced by Wightman functions<sup>4</sup>

$$\langle \mathcal{O}_1 \zeta(x_1) \mathcal{O}_2 \zeta(x_2) \mathcal{O}_3 \rangle_0 \rightarrow \langle \zeta(x_1) \zeta(x_2) \rangle_0 \langle \mathcal{O}_1 \mathcal{O}_2 \mathcal{O}_3 \rangle_0 + \dots \quad (3.7)$$

Tadpoles can be removed by normal ordering the interacting Hamiltonian i.e. in practice neglecting self contractions. The choice of vacuum is then equivalent to the choice of Wightman functions

$$\langle \zeta(x_1) \zeta(x_2) \rangle_0 = \int \frac{d^3 k}{(2\pi)^3} \mathcal{V}_k(\eta_1) \mathcal{V}_k^*(\eta_2) e^{i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)}. \quad (3.8)$$

The modes  $\mathcal{V}_k$  satisfy the free field equation

$$\mathcal{V}_k'' + 2\frac{z'}{z}\mathcal{V}_k + k^2\mathcal{V}_k = 0, \quad (3.9)$$

where  $z = a\dot{\phi}/H$ . We choose them to be normalised as

$$\mathcal{V}_k \rightarrow \frac{1}{\sqrt{2kz}} (\alpha_k e^{-ik\eta} + \beta_k e^{ik\eta}), \quad (3.10)$$

in the limit  $|k\eta| \gg 1$ . The three point function will then be given by a causal time integral over products of Wightman functions and their derivatives. We are now ready to turn to the specific calculations.

### 3.1 Scalar Minimally Coupled to Gravity

Consider now the situation treated by Maldacena, where the inflaton is a scalar field minimally coupled to gravity. The result contains two contributions, one from a local (in time) nonlinear field redefinition, and the second from an integral of the form in Eq. (3.5). Since local redefinitions will not contain the enhancement terms we are interested in, we can concentrate on the non-local contribution. Maldacena [7] showed that up to these redefinitions the action contains the following three point interaction for  $\zeta$  (we have reexpressed his result in conformal time)

$$S_{(3)} = \int d\eta d^3x a^3 M_{\text{pl}}^{-2} \left( \frac{\dot{\phi}}{H} \right)^4 H \zeta'^2 \partial^{-2} \zeta', \quad (3.11)$$

where  $\dot{\phi}$  denotes the *cosmic* time derivative of the inflaton zero mode. We use it instead of the conformal time derivative since the slow-roll conditions are easier to

---

<sup>4</sup>This is the “in-in” version of the normal Wick's theorem for time ordered products.

express in terms of  $\dot{\phi}$ ;  $\dot{\phi} \simeq \sqrt{2\epsilon} M_{\text{pl}} H$ . The associated interacting Hamiltonian is  $H_I = - \int d^3x a^3 \left( \frac{\dot{\phi}}{H} \right)^4 H \zeta'^2 \partial^{-2} \zeta'$ . Note that this conversion is not entirely trivial due to the time-derivative dependence of the interaction. Substituting in the general expression Eq. (3.6) we obtain

$$\begin{aligned} \langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle &= -i(2\pi)^3 \delta^3(\sum \vec{k}_i) \left( \frac{\dot{\phi}}{H} \right)^4 M_{\text{pl}}^{-2} H \int_{\eta_0}^0 d\eta a^3(\eta) \frac{1}{k_3^2} \partial_\eta G_{k_1}^>(0, \eta) \partial_\eta G_{k_2}^>(0, \eta) \partial_\eta G_{k_3}^>(0, \eta) \\ &+ \text{permutations} + \text{c.c.} . \end{aligned} \quad (3.12)$$

In writing the above expression, we have evaluated the three point function at  $\eta = 0$ , i.e. when the modes are well outside the Hubble-horizon. Here we see explicitly how the product of Wightman functions appears. For the BD vacuum the Wightman function is given by

$$G_{k_1}^>(\eta, \eta') = \frac{H^2}{\dot{\phi}^2} \frac{H^2}{2k^3} (1 + ik\eta)(1 - ik\eta') e^{-ik(\eta - \eta')}. \quad (3.13)$$

For the first argument taken to be well after horizon crossing, we have

$$\partial_\eta G_{k_1}^>(0, \eta) = -\frac{H}{2k} \left( \frac{H}{\dot{\phi}} \right)^2 \frac{1}{a(\eta)} e^{ik\eta} \quad (3.14)$$

and so (in the limit  $\eta_0 \rightarrow -\infty$ )

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = (2\pi)^3 \delta^3(\sum \vec{k}_i) \frac{1}{\prod (2k_i^3)} \frac{4H^6}{M_{\text{pl}}^2 \dot{\phi}^2} \frac{\sum_{i>j} k_i^2 k_j^2}{k_t}, \quad (3.15)$$

where  $k_t = k_1 + k_2 + k_3$ , and  $k_i = |\vec{k}_i|$ .

Suppose we now modify the mode functions:  $\mathcal{U}_k \rightarrow \mathcal{V}_k$ . We get two types of corrections: one from the modifications to the norm of the positive frequency modes  $\alpha_k \mathcal{U}_{k_i}$  which just changes the overall normalisation and shape dependences by an amount similar to that for the power spectrum. The second type of corrections which are suppressed by powers of  $\beta_k$  are more interesting. One such correction will again modify the over amplitude, but the more relevant corrections are those that change the arguments of the oscillatory exponentials. For instance, to linear order in  $\beta_k$  we have the correction

$$\Delta \langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = -i(2\pi)^3 \delta^3(\sum \vec{k}_i) \frac{2}{\prod (2k_i^3)} \frac{H^6}{M_{\text{pl}}^2 \dot{\phi}^2} \int_{\eta_0}^0 d\eta \sum_j \beta_{k_j}^* \frac{k_1^2 k_2^2 k_3^2}{k_j^2} e^{i\tilde{k}_j \eta} + \text{c.c.}, \quad (3.16)$$

where we have defined  $\tilde{k}_j = k_t - 2k_j$ . Performing the integral we have

$$\Delta \langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = -(2\pi)^3 \delta^3(\sum \vec{k}_i) \frac{2}{\prod (2k_i^3)} \frac{H^6}{M_{\text{pl}}^2 \dot{\phi}^2} \sum_j \frac{k_1^2 k_2^2 k_3^2}{k_j^2 \tilde{k}_j} \beta_{k_j}^* \left( 1 - e^{i\tilde{k}_j \eta_0} \right) + \text{c.c.} . \quad (3.17)$$

What makes this contribution interesting are the denominators proportional to  $\tilde{k}_j$  which appear. In the standard calculation in Ref. [7], the denominator only involved  $k_t = k_1 + k_2 + k_3$  which could never vanish unless all the  $k$ 's vanished, which would then force the numerator to vanish as well. However, the mixing between positive and negative energy BD states in our modes now allows for vanishing denominators ( $\tilde{k}_j = 0$ ) for certain *non-zero* values of the momenta, which *enhances* these contributions relative to those for the BD vacuum. In practice there is no divergence since the exponential factor is unity in the  $\tilde{k}_j \rightarrow 0$  limit. However, this cutoff is only relevant for  $\tilde{k}_j \approx 1/\eta_0$ . Thus if we evaluate Eq. (3.17) for these special triangles where  $\tilde{k}_j = 0$ , we get

$$\Delta\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle|_{\tilde{k}_j=0} = (2\pi)^3\delta^3(\sum\vec{k}_i)\frac{4}{\prod(2k_i^3)}\frac{H^6}{M_{\text{pl}}^2\phi^2}\frac{k_1^2k_2^2k_3^2}{k_j^2}\text{Im}(\beta_{k_j})\eta_0, \quad (3.18)$$

and so relative to the standard result we get an enhancement factor

$$\frac{\Delta\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle}{\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle}\bigg|_{\tilde{k}_j=0} = \frac{\frac{k_1^2k_2^2k_3^2}{k_j^2}\text{Im}(\beta_{k_j})k_t\eta_0}{\sum_s\frac{k_1^2k_2^2k_3^2}{k_s^2}} \approx |\beta_k||k\eta_0| = |\beta_k|\frac{k}{a(\eta_0)H}. \quad (3.19)$$

We have implicitly assumed that  $\text{Im}\beta_k \sim |\beta_k|$ . Note then that although the result is suppressed by a factor of  $\beta_k$ , this is multiplied by the ratio of the physical momentum at the beginning of inflation to the Hubble scale, which may be as large as  $M/H$  (at least for those modes which have been excited). In practice however, we cannot measure the full three point function but only the two dimensional projection of it encoded into the CMB. As a result we must effectively average this result over angles which we will do in Sec. 4. On doing so we essentially lose a factor of  $|k\eta_0|$  and so this result is not enhanced but will be of the same order as similar small corrections from the local redefinitions and to the two-point function.

Nevertheless, for interactions which scale with larger powers of  $1/a$  the enhancement effect can be significant and we shall discuss these in the next section. The enhancement will occur when *e.g.*  $k_1 + k_2 = k_3$ . From the momentum delta function, we also have that  $|\vec{k}_1 + \vec{k}_2| = \vec{k}_3$  so that only “flattened” triangles with two of the vectors being collinear will be enhanced. Whether or not the effect is observable then has to do with whether these triangles have sufficient “measure” when the bi-spectrum is converted to  $\ell$ -space. We will discuss this further below.

### 3.2 Higher derivative interactions

Now let us turn to the model of higher derivative interactions discussed in [6]. It contains the following dimension 8 correction to the effective action for the inflaton:

$$\mathcal{L}_I = \sqrt{-g}\frac{\lambda}{8M^4}((\nabla\Phi)^2)^2. \quad (3.20)$$

This type of correction is regularly considered in modifications to the kinetic term of the inflaton such as in k-inflation [33], or DBI inflation [34]. Since we do not want to be wedded to a specific scenario, for the purposes of the present discussion we shall concentrate on standard slow-roll inflation; however, it is straightforward to extend the arguments to these more general cases.

Before looking at its contribution to the three point function, we shall make some cautionary remarks about the tree level contribution to the two point function coming from this interaction. Expanding  $\Phi = \phi(\eta) + \delta\phi$  to second order in  $\delta\phi$  and then converting from  $\delta\phi$  to  $\zeta$ , this term contributes the following term to the Hamiltonian

$$\begin{aligned} H_I^{(2)} &= \frac{\lambda}{4M^4} \int d^3x a^2 \dot{\phi}^2 \left( 3\delta\phi'^2 - (\partial_i \delta\phi)^2 \right) \\ &= \frac{\lambda}{4M^4} \int d^3x a^2 \frac{\dot{\phi}^4}{H^2} \left( 3(\zeta' + ar\zeta)^2 - (\partial_i \zeta)^2 \right), \end{aligned} \quad (3.21)$$

where  $r = \frac{d \ln(\dot{\phi}/H)}{dt}$ . Since  $r$  is both slow roll suppressed and its contribution to the action subdominant at subhorizon scales we can ignore it for the present discussion. Since this term is quadratic in the fields, we could, in principle, simply incorporate it into the free Hamiltonian. On the other hand, we could also treat this as an interaction, just as for the three point function. Doing this would yield the following correction to the  $\zeta$  equal time two point function evaluated at late times

$$\Delta \langle \zeta(\vec{x}_1) \zeta(\vec{x}_2) \rangle = -\frac{i\lambda}{2M^4} \int_{\eta_0}^0 d\eta a^2(\eta) \frac{\dot{\phi}^4}{H^2} \left( 3(\partial_\eta G_k^>(0, \eta))^2 - k^2 G_k^>(0, \eta)^2 \right) + \text{c.c.} \quad (3.22)$$

On considering the terms linear in  $\beta_k$ , this appears to give similar enhancements as described in the previous section. In this case the enhancements will not get washed out by averaging over angles. However, this is clearly a fake since had we incorporated this term into the free Hamiltonian we would not have seen any such affect. The resolution is that this term can be removed by a renormalization of the initial vacuum choice  $\beta_k \rightarrow \beta_k + \lambda \delta \beta_k$ .

To see where the problem arises we can always resum two point vertices by absorbing them into the definition of the free Hamiltonian. On doing so, the Wightman function will be constructed from free modes satisfying an equation of the form

$$\mathcal{V}_k'' + 2\frac{z'}{z}\mathcal{V}_k + c_s^2 k^2 \mathcal{V}_k = 0, \quad (3.23)$$

where

$$c_s^2 = \frac{(1 - \lambda \dot{\phi}^2/(2M^2))}{(1 - \lambda \dot{\phi}^2/(3M^2))}. \quad (3.24)$$



Here we see the modified sound speed characteristic of models such as k-inflation [33] and DBI inflation [34] that make use of these higher derivative operators. At subhorizon scales the WKB approximation is valid, and so can write the general solution as

$$\mathcal{V}_k = \frac{1}{\sqrt{2c_s k z}} (\alpha_k \exp(-ic_s |k| \eta) + \beta_k \exp(+ic_s |k| \eta)). \quad (3.25)$$

We can now see the reason for the apparently large corrections to the power spectrum. If we expand these modes out to  $\mathcal{O}(\lambda)$  we get

$$\exp(-ic_s |k| \eta) = \exp(-i|k| \eta) (1 - i|k| \eta \frac{\lambda \dot{\phi}^2}{2M^4} + \mathcal{O}(\lambda^2)), \quad (3.26)$$

which is a poor expansion for sufficiently large  $|k\eta|$ . From now on we shall work with the resummed mode functions  $\mathcal{V}_k$  which are well behaved at subhorizon scales, which has the effect of removing the correction in Eq. (3.22); since the correction to the three point function computed below is already  $\mathcal{O}(\lambda)$ , we can omit these corrections to the modes.

Having taking care of these issues, we can now turn to the calculation of the three point function in the presence of these higher derivative interactions. Expanding the interaction Lagrangian to third order in the fluctuations  $\delta\phi$  and then converting to  $\zeta$ , we find

$$\begin{aligned} H_I^{(3)} &= - \int d^3 x a \frac{\lambda \dot{\phi}}{2M^4} \delta\phi' \left( \delta\phi'^2 - (\partial_i \delta\phi)^2 \right), \\ &= - \int d^3 x a \frac{\lambda \dot{\phi}^4}{2H^3 M^4} \zeta' \left( \zeta'^2 - (\partial_i \zeta)^2 \right), \end{aligned} \quad (3.27)$$

where again we have neglected the  $r$  term. Note that although the original operator was dimension 8, this operator is really dimension 6 in terms of  $\delta\phi$ , but it is additionally suppressed by  $\sqrt{\epsilon} (HM_{\text{pl}}/M^4)$ .

We can now compute the corrections of interest to the three point function. We compute to lowest order in  $\lambda$ ,  $H/M$ , and  $\epsilon$  to find

$$\begin{aligned} \langle \zeta_{k_1} \zeta_{k_2} \zeta_{k_3} \rangle &= (2\pi)^3 \delta^{(3)} \left( \sum_i \vec{k}_i \right) \mathcal{A}(\vec{k}_1, \vec{k}_2, \vec{k}_3) \\ \mathcal{A}(\vec{k}_1, \vec{k}_2, \vec{k}_3) &= i \frac{\lambda \dot{\phi}^4}{2H^3 M^4} \int_{\eta_0}^0 d\eta a \left( \partial_\eta G_{k_1}^>(0, \eta) \partial_\eta G_{k_2}^>(0, \eta) \partial_\eta G_{k_3}^>(0, \eta) + \right. \\ &\quad \left. + (\vec{k}_1 \cdot \vec{k}_2) (G_{k_1}^>(0, \eta) G_{k_2}^>(0, \eta) \partial_\eta G_{k_3}^>(0, \eta)) + \text{perms.} \right) + \text{c.c.} \end{aligned} \quad (3.28)$$

The corrections we want are again those linear in  $\beta_k$ . These give rise to the following contribution to the three point function  $\zeta_k$ .

$$\Delta\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_3}\zeta_{\vec{k}_3}\rangle=(2\pi)^3\delta^{(3)}\left(\sum_i\vec{k}_i\right)\frac{1}{\prod(2k_i^3)}\frac{i\lambda H^8}{\dot{\phi}^2 M^4}\sum_j\beta_{k_j}^*\int_{\eta_0}^0 d\eta e^{i\vec{k}_j\eta}S_j(k_1,k_2,k_3,\eta)+\text{c.c.}, \quad (3.29)$$

where

$$S_j(k_1,k_2,k_3,\eta)=-k_t\left(\Pi_i\tilde{k}_i\right)+ \quad (3.30)$$

$$(-\eta)\left(k_j^4(k_{j+1}+k_{j+2})+k_{j+1}k_{j+2}(k_{j+1}-k_{j+2})^2(k_{j+1}+k_{j+2})-k_j^3(k_{j+1}^2+k_{j+2}^2)\right. \\ \left.-k_j^2(k_{j+1}+k_{j+2})(k_{j+1}^2+k_{j+1}k_{j+2}+k_{j+2}^2)\right)+(-\eta)^2\tilde{k}_j\left(\Pi_i k_i\right)\left(k_t^2-4k_{j+1}k_{j+2}\right),$$

where  $j$  is defined modulo 3. As in Sec. 3.1, we see that these results are enhanced for the flattened triangles where two of the vectors are collinear so that  $\tilde{k}_i$  vanishes for some  $i$ . Note that the first and third terms in the  $\eta$  expansion of  $S_j$  vanish and so the dominant contribution evaluated on a give flattened triangle  $\tilde{k}_j=0$  is

$$\Delta\langle\zeta_{\vec{k}_1}\zeta_{\vec{k}_3}\zeta_{\vec{k}_3}\rangle|_{\tilde{k}_j=0}=-(2\pi)^3\delta^{(3)}\left(\sum_i\vec{k}_i\right)\frac{1}{\prod(2k_i^3)}\frac{\lambda H^8}{\dot{\phi}^2 M^4}2\mathcal{R}e(\beta_{k_j}) \quad (3.31)$$

$$\times\frac{\eta_0^2}{2}\left(4k_{j+1}k_{j+2}(k_{j+1}+k_{j+2})(k_{j+1}^2+k_{j+2}^2+k_{j+1}k_{j+2})\right).$$

For these specific triangles the enhancement factor relative to the BD contribution is

$$\frac{\Delta\langle\zeta^3\rangle}{\langle\zeta^3\rangle}\Big|_{\text{flattened}}\approx|\beta_k||k\eta_0|^2, \quad (3.32)$$

where we have assumed  $\mathcal{R}e(\beta_k)\approx|\beta_k|$ . Again as will be explained in the next section on going to  $l$  space we effectively lose one factor of  $|k\eta_0|$ , but one enhancement factor remains. Thus we can give an order of magnitude estimate for  $f_{NL}$  for these triangles to be

$$f_{NL}|_{\text{flattened}}\sim\frac{\dot{\phi}^2}{M^4}|\beta_k|\left(\frac{k}{a(\eta_0)H}\right). \quad (3.33)$$

For backreaction to be under control, the largest reasonable value for  $|\beta_k|$  is  $\sqrt{\epsilon\eta'}HM_{\text{pl}}/M^2$  while for the effective field theory to be valid the largest value allowed for  $k/a(\eta_0)$  is  $M$ . Thus the maximum expected contribution to  $f_{NL}$  is

$$f_{NL}|_{\text{flattened}}\sim\sqrt{\epsilon\eta'}\frac{\dot{\phi}^2}{M^4}\frac{HM_{\text{pl}}}{M^2}\left(\frac{M}{H}\right)\sim\epsilon\sqrt{\epsilon\eta'}\left(\frac{H}{M_{\text{pl}}}\right)^2\left(\frac{M_{\text{pl}}}{M}\right)^5. \quad (3.34)$$

We see then that for reasonable values of these parameters  $\epsilon, \eta' \sim 10^{-2}$ ,  $H/M_{\text{pl}} \sim 10^{-6}$  we get

$$f_{NL}|_{\text{flattened}} \sim \left(6.3 \times 10^{-4} \frac{M_{\text{pl}}}{M}\right)^5, \quad (3.35)$$

which is in the range 1 to 100 for  $M$  in the range  $6.3 \times 10^{-4} M_{\text{pl}}$  to  $2.5 \times 10^{-4} M_{\text{pl}}$ . In pushing into the limit  $f_{NL} \sim 100$  we have  $|\beta_k| \sim 1.6 \times 10^{-1}$  which may or may not be already observable in the power spectrum depending on the precise  $k$  dependence of the  $\beta_k$ . In practice these bounds are overly restrictive since we know that models such as DBI inflation do not have to satisfy the usual slow roll restrictions, and in fact the more general case was considered in Ref. [17]. In these models the usual contribution to the non-gaussianity may be large, and so this specific contribution will be further enhanced. In general the enhancement factor for these specific triangles relative to the usual contribution is bounded by

$$\frac{f_{NL}|_{\text{flattened}}}{f_{NL}|_{\text{usual}}} \sim |\beta_k| \frac{M}{H} \sim 100 \sqrt{|\beta_k|}. \quad (3.36)$$

From an observational point of view, one might think that since only specific triangles give the enhancement, this effect might be buried in the noise. However, consider the  $l = 500$  total modes which have been observed by WMAP with signal to noise greater than 1. In total, the bi-spectrum is made of  $(500)^3$  points (modulo symmetry factors). Of these, the total number that satisfy the triangle inequality are roughly  $(500)^2$  and so are down by a factor of 500. This would then imply that the signal to noise coming from these will thus be down by a factor of  $1/\sqrt{500}$ . Thus, roughly 5 percent of the signal will come from these modes, and so to a crude approximation  $\sim 10^2 \sqrt{|\beta_k|}$  enhancement will effectively only amount to a  $\sim 10 \sqrt{|\beta_k|}$  enhancement. In practice it is necessary to reanalyse the data with an appropriate template along the lines of Ref. [30], and this will improve the constraints on this contribution.

### 3.3 Backreaction from higher derivative interactions

We have already given a simple estimate for the absence of backreaction based on computing the expectation value of the free field stress energy. However, as soon as we choose a non-standard initial state, and a non-zero gaussianity develops we will also generate additional cubic and higher order contributions to the stress energy. As usual we can infer these via the in-in formalism [23] so that the expectation value of the stress energy at any time  $\eta$  is given by

$$\langle T_{\mu\nu}(\eta, x) \rangle = \langle \bar{T} e^{i \int_{\eta_0}^{\eta} d\eta' H_I} T_{\mu\nu}(\eta, x) T e^{-i \int_{\eta_0}^{\eta} d\eta' H_I} \rangle_0, \quad (3.37)$$

which gives to first order in the interaction

$$\langle T_{\mu\nu}(\eta, x) \rangle = -2\mathcal{R}e \left( i \int_{\eta_0}^{\eta} d\eta' \langle T_{\mu\nu}(\eta, x) H_I \rangle_0 \right), \quad (3.38)$$

where it is important to not that the  $T_{\mu\nu}$  inserted on the right hand side is defined in the interacting representation. Thus in particular a contribution to  $T_{\mu\nu}$  cubic in the interaction multiplied by the cubic vertex in the Hamiltonian will give a non-zero expectation value. If we choose the Bunch-Davies vacuum then for the usual reasons we can safely ignore this backreaction contribution since the vacuum is de Sitter invariant and it will at most contribute to a renormalization of the zero point of the potential. If, however, we choose instead an excited yet still gaussian initial state, the early non-gaussianity that is generated may cause a problem.

Adding to the action the interaction  $\frac{\lambda}{8M^4}(\partial\phi)^4$  gives the following contribution to the stress energy (in the Heisenberg rep.)

$$\Delta T_{\mu\nu}^H = -\frac{\lambda}{2M^4}(\partial\phi)^2 \partial_\mu \phi \partial_\nu \phi + \frac{\lambda}{8M^4}(\partial\phi)^4 g_{\mu\nu}. \quad (3.39)$$

In particular, expanding to cubic order around an inflating solution  $\phi(\eta)$  (and neglecting backreaction of the metric which is a reasonable approximation at sub-Hubble scales) we get the following cubic contribution to the energy density

$$\Delta_{(3)}\rho^H = \frac{\lambda\dot{\phi}}{2M^4 a^3} \delta\phi' \left( 3\delta\phi'^2 - (\partial_i \delta\phi)^2 \right). \quad (3.40)$$

On converting to the interacting representation we unsurprisingly obtain the interacting Hamiltonian density<sup>5</sup>

$$\Delta_{(3)}\rho = \frac{\lambda\dot{\phi}}{2M^4 a^3} \delta\phi' \left( \delta\phi'^2 - (\partial_i \delta\phi)^2 \right). \quad (3.41)$$

In any case rather than performing the full calculation we may simply note that at sub-Hubble scales  $\delta\phi'$  and  $\partial_i \delta\phi$  scale the same way so the effective contributions to the pressure and stress energy take the form (where we have defined  $\chi = a\delta\phi$ )<sup>6</sup>

$$\Delta_{(3)}\rho \approx \frac{\lambda^2 \dot{\phi}^2}{M^8 a^6(\eta)} \int_{\eta_0}^{\eta} d\eta' \int d^3x' \frac{1}{a^2(\eta')} \langle \chi'^3(\eta, x) \chi'^3(\eta', x') \rangle_0 \quad (3.42)$$

$$\approx \frac{\lambda^2 \dot{\phi}^2}{M^8 a^6(\eta)} \int_{\eta_0}^{\eta} d\eta' \int d^3x' \frac{1}{a^2(\eta')} (\partial_\eta \partial_{\eta'} H(x', \eta - \eta'))^3, \quad (3.43)$$

---

<sup>5</sup>To get from one to the other one first writes the Heisenberg expression in terms of  $\phi$  and  $\pi$ , and then replace  $\pi$  by its free field value  $\pi = a^3 \dot{\phi}$ .

<sup>6</sup>In fact the final answer is smaller since the Lorentz invariant combination  $\delta\phi'^2 - (\partial_i \delta\phi)^2$  gives an additional  $H/M$  suppression.

where

$$H(x', \eta - \eta') = \int \frac{d^3 k}{(2\pi)^3} \frac{1}{2|k|} e^{-i\vec{k} \cdot \vec{x}'} \left( |\alpha_k|^2 e^{-i|k|(\eta - \eta')} + \alpha_k^* \beta_k e^{i|k|(\eta + \eta')} + \alpha_k \beta_k^* e^{-i|k|(\eta + \eta')} + |\beta_k|^2 e^{i|k|(\eta - \eta')} \right). \quad (3.44)$$

This gives at worst a contribution of order

$$\Delta_{(3)} \rho \approx \frac{\lambda^2 \dot{\phi}^2}{M^8 a^6} \mathcal{R}e \int_{\eta_0}^{\eta} d\eta' \frac{1}{a^2(\eta')} \int \frac{d^3 k_1}{(2\pi)^3} \int \frac{d^3 k_2}{(2\pi)^3} k_1 k_2 |\vec{k}_1 + \vec{k}_2| |\beta_{k_1}^*| e^{-2ik_1 \eta} e^{-i(-k_1 + k_2 + |\vec{k}_1 + \vec{k}_2|)(\eta - \eta')}, \quad (3.45)$$

to first order in  $\beta_k$ . As explained in Sec. 2 the presence of the oscillatory factor  $e^{\pm 2ik_1 \eta}$  washes out the early time contribution to this integral. The leading contribution is then the  $O(\beta_k^2)$  contribution which is bounded by

$$\Delta \rho_{(3)} \lesssim |\beta_0|^2 \dot{\phi}^2. \quad (3.46)$$

To obtain this estimate we can simply count the positive powers of  $k$  and  $\eta$  in the integral and replace them with their maximum possible values  $a(\eta_0)M$  and  $\eta_0$  respectively, and then dividing by one factor of  $|k\eta_0|$  from the angular integration. Note that since we also require  $\dot{\phi}^2 < M^4$  for the validity of the effective field theory this contribution is necessarily smaller than the free field contribution.

Similarly at quartic order we have in the Heisenberg representation

$$\Delta_{(4)} \rho^H = \frac{\lambda}{8M^4 a^4} \left( 3\delta\phi'^4 - 2(\delta\phi')^2 (\partial_i \delta\phi)^2 - (\partial_i \delta\phi)^4 \right). \quad (3.47)$$

This gives a contribution of the form

$$\Delta_{(4)} \rho \approx \frac{\lambda^2}{M^8 a^8(\eta)} \int_{\eta_0}^{\eta} d\eta' \int d^3 x' \frac{1}{a^4(\eta')} \langle \chi'^4(\eta, x) \chi'^4(\eta', x') \rangle_0 \quad (3.48)$$

$$\approx \frac{\lambda^2}{M^8 a^8(\eta)} \int_{\eta_0}^{\eta} d\eta' \int d^3 x' \frac{1}{a^4(\eta')} (\partial_\eta \partial_{\eta'} H(x', \eta - \eta'))^4. \quad (3.49)$$

As before the dominant contribution is bounded by

$$\Delta \rho_{(4)} \lesssim |\beta_0|^2 M^4. \quad (3.50)$$

According to the general argument given in Sec. 2 higher order contributions will never give a contribution larger than this as long as we remain in the validity of the effective theory.

## 4. Measure issues and the l-space bi-spectrum

The results of the previous section show that for flattened triangles the three point function can be considerably enhanced relative to the usual case. However, we do not actually measure the three dimensional three point function directly in the CMB. Rather, we measure its two dimensional projection as encoded in the  $a_{lm}$ 's.

The relevant question then is to what extent these enhancements can be seen in the CMB l-space three point function. Once the temperature fluctuations have been decomposed into spherical harmonics,

$$\frac{\Delta T}{T}(\hat{n}) = \sum_{lm} a_{lm} Y_{lm}(\hat{n}), \quad (4.1)$$

we can construct the angular averaged bi-spectrum [35]

$$B(l_1, l_2, l_3) = \sum_{m_1, m_2, m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle. \quad (4.2)$$

Expressing the three point function for  $\zeta_k$  as

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = (2\pi)^3 \delta\left(\sum_i \vec{k}_i\right) \mathcal{A}(\vec{k}_1, \vec{k}_2, \vec{k}_3), \quad (4.3)$$

the three point function of the  $a_{lm}$  is given by (in this section we will make use of the unit vectors defined via  $\vec{k}_i = |k_i| \hat{n}_i$ ).

$$\begin{aligned} \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle &= (4\pi)^3 i^{l_1+l_2+l_3} \int \frac{d^3 \vec{k}_1}{(2\pi)^3} \frac{d^3 \vec{k}_2}{(2\pi)^3} \frac{d^3 \vec{k}_3}{(2\pi)^3} Y_{l_1 m_1}^*(\hat{n}_1) Y_{l_2 m_2}^*(\hat{n}_2) Y_{l_3 m_3}^*(\hat{n}_3) \\ &\quad (2\pi)^3 \delta^{(3)}\left(\sum_i \vec{k}_i\right) \mathcal{A}(\vec{k}_1, \vec{k}_2, \vec{k}_3) \Delta_{l_1}^T(k_1) \Delta_{l_2}^T(k_2) \Delta_{l_3}^T(k_3), \end{aligned} \quad (4.4)$$

where  $\Delta_l^T(k)$  are the radiation transfer functions. We are interested in a specific combination to  $\mathcal{A}$  that comes from flattened triangles. The contributions to the three point function linear in  $\beta_k$  can be written in the form

$$\mathcal{A}(\vec{k}_1, \vec{k}_2, \vec{k}_3) = \sum_j \beta_{k_j}^* (i k_t)^{s+1} \int_{\eta_0}^0 d\eta (-\eta)^s e^{i \vec{k}_j \eta} \mathcal{D}(k_j, k_{j+1}, k_{j+2}) + c.c., \quad (4.5)$$

where  $j$  is defined modulo 3. Let us consider the properties of the integral

$$I(\tilde{k}_i) = \int_{\eta_0}^0 d\eta (-\eta)^s e^{i \tilde{k}_i \eta}. \quad (4.6)$$

For  $|\tilde{k}_i \eta_0| \gg 1$  the oscillatory nature of the integrand heavily damps the power of  $\eta$  and as a result the integral scales as  $(1/i\tilde{k}_i)^{s+1}$ . On the other hand for  $|\tilde{k}_i \eta_0| \ll 1$  the exponential is irrelevant and the integral gives  $I(\tilde{k}_i) = (-\eta_0)^{s+1}/(s+1)$ . We are interested in effects coming from sub-Horizon scales where  $|k_i \eta_0| \gg 1$ , and so the dominant contribution to the integral will come from the regime where the triangles are flattened and  $\tilde{k}_i \approx 0$  so that the integral is given by its maximum value  $I(\tilde{k}_i) = (-\eta_0)^{s+1}/(s+1)$ . However, this will only occur for a special range of angles, essentially for

$$\hat{n}_i \cdot \hat{n}_j + 1 \ll \frac{(k_i - k_j)^2}{k_i k_j}. \quad (4.7)$$

To perform the integrals, let us first integrate out  $\vec{k}_{j+2}$  via the delta function, and then integrate over  $\hat{n}_i$  using  $\hat{n}_{i+1}$  as a reference axis. On doing so we get

$$\begin{aligned} \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle &= \sum_j (4\pi)^3 i^{l_1+l_2+l_3} \int \frac{dk_j k_j^2}{(2\pi)^3} \int \frac{dk_{j+1} k_{j+1}^2}{(2\pi)^3} \beta_j^* \int d^2 \hat{n}_{j+1} \int d^2 \hat{n}_j \\ &Y_{l_j m_j}^*(\hat{n}_j) Y_{l_{j+1} m_{j+1}}^*(\hat{n}_{j+1}) Y_{l_{j+2} m_{j+2}}^*(\hat{n}_{j+2}) (ik_t)^{s+1} I(\tilde{k}_j) \mathcal{D}(k_j, k_{j+1}, |\vec{k}_j - \vec{k}_{j+1}|) \times \\ &\Delta_{l_j}^T(k_j) \Delta_{l_{j+1}}^T(k_{j+1}) \Delta_{l_{j+2}}^T(|\vec{k}_j - \vec{k}_{j+1}|) + \dots, \end{aligned} \quad (4.8)$$

where the ellipsis represents the terms arising from the complex conjugate in Eq. (4.5). At this point we can make the approximation that since the integral  $I(\tilde{k}_j)$  is so strongly peaked at  $\tilde{k}_j = 0$  we may replace the product of spherical harmonics with their values for the flattened triangles set by  $\tilde{k}_j = 0$ , namely

$$Y_{l_j m_j}^*(\hat{n}_j) Y_{l_{j+1} m_{j+1}}^*(\hat{n}_{j+1}) Y_{l_{j+2} m_{j+2}}^*(\hat{n}_{j+2}) \rightarrow Y_{l_j m_j}^*(-\hat{n}_{j+1}) Y_{l_{j+1} m_{j+1}}^*(\hat{n}_{j+1}) Y_{l_{j+2} m_{j+2}}^*(\hat{n}_{j+1}). \quad (4.9)$$

Similarly we may replace

$$\Delta_{l_{j+2}}^T(|\vec{k}_j - \vec{k}_{j+1}|) \rightarrow \Delta_{l_{j+2}}^T(|k_j - k_{j+1}|). \quad (4.10)$$

In this approximation, the integral  $I(\tilde{k}_j)$  gets directly integrated over angles

$$\int d^2 \hat{n}_j I(\tilde{k}_j) = 2\pi \int_{\eta_0}^0 d\eta (-\eta)^s \frac{1}{k_j k_{j+1} \eta^2} \left( e^{2ik_{j+1}\eta} (1 - i(k_j + k_{j+1})\eta) - (1 - i(k_j - k_{j+1})\eta) \right), \quad (4.11)$$

where we have assumed  $k_j \geq k_{j+1}$  as is necessary for the existence of this triangle. In the limit  $|k_s \eta_0| \gg 1$  this integral is given approximately by

$$\int d^2 \hat{n}_j I(\tilde{k}_j) \approx \frac{2\pi i (k_j - k_{j+1})}{s k_j k_{j+1}} (-\eta_0)^s. \quad (4.12)$$

So we have

$$\begin{aligned}
\langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle &\approx \sum_j (4\pi)^3 i^{l_1+l_2+l_3} \int \frac{dk_j k_j^2}{(2\pi)^3} \int_{k_{j+1} < k_j} \frac{dk_{j+1} k_{j+1}^2}{(2\pi)^3} \int d^2 \hat{n}_{j+1} \\
&Y_{l_j m_j}^*(-\hat{n}_{j+1}) Y_{l_{j+1} m_{j+1}}^*(\hat{n}_{j+1}) Y_{l_{j+2} m_{j+2}}^*(\hat{n}_{j+1}) \beta_{k_j}^* \frac{2\pi k_t (k_{j+1} - k_j)}{s k_j k_{j+1}} (-i k_t \eta_0)^s \times \\
&\mathcal{D}(k_j, k_{j+1}, |k_j - k_{j+1}|) \Delta_{l_j}^T(k_j) \Delta_{l_{j+1}}^T(k_{j+1}) \Delta_{l_{j+2}}^T(|k_j - k_{j+1}|) + \dots
\end{aligned} \tag{4.13}$$

Using the fact that  $Y_{lm}(-\hat{n}) = (-1)^l Y_{lm}(\hat{n})$  we get

$$\begin{aligned}
\langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle &\approx \sum_j (4\pi)^3 i^{l_1+l_2+l_3} (-1)^{l_1} \int \frac{dk_j k_j^2}{(2\pi)^3} \int_{k_{j+1} < k_j} \frac{dk_{j+1} k_{j+1}^2}{(2\pi)^3} \beta_{k_j}^* \mathcal{G}_{l_1, l_2, l_3}^{m_1, m_2, m_3} \\
&\frac{2\pi k_t (k_{j+1} - k_j)}{s k_j k_{j+1}} (-i k_t \eta_0)^s \mathcal{D}(k_j, k_{j+1}, |k_j - k_{j+1}|) \Delta_{l_1}^T(k_j) \Delta_{l_2}^T(k_{j+1}) \Delta_{l_3}^T(|k_j - k_{j+1}|) + \dots,
\end{aligned}$$

where  $\mathcal{G}_{l_1, l_2, l_3}^{m_1, m_2, m_3}$  is the Gaunt integral [35]. Finally then, these contributions to the bi-spectrum take the form

$$\begin{aligned}
B(l_1, l_2, l_3) &= \\
&\sum_j \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} (4\pi)^3 i^{l_1+l_2+l_3} (-1)^{l_1} \int \frac{dk_j k_j^2}{(2\pi)^3} \int_{k_{j+1} < k_j} \frac{dk_{j+1} k_{j+1}^2}{(2\pi)^3} \\
&\beta_{k_j}^* \frac{2\pi k_t (k_{j+1} - k_j)}{s k_j k_{j+1}} (-i k_t \eta_0)^s \mathcal{D}(k_j, k_{j+1}, |k_j - k_{j+1}|) \Delta_{l_j}^T(k_j) \Delta_{l_{j+1}}^T(k_{j+1}) \Delta_{l_{j+2}}^T(|k_j - k_{j+1}|) + \text{c.c.}
\end{aligned} \tag{4.14}$$

The crucial point to note is that whilst the original integral naively scales as  $|k_t \eta_0|^{s+1}$ , the above bi-spectrum only contains a factor  $|k_t \eta_0|^s$ . In short, in performing the 2-d projection we essentially have to smooth the triangles over some finite resolution of solid angles. This softens the size of the enhancement by one power of  $|k \eta_0| = k/(a(\eta_0)H)$ . Nevertheless as we have described in Sec. 3 the remaining effect can feasibly be larger than the usual one. The radiation transfer functions are dominated by the multipole  $l \approx k d_{LSS}$  where  $d_{LSS}$  is the distance to the last scattering surface, and so it is clear that the product  $\Delta_{l_j}^T(k_j) \Delta_{l_{j+1}}^T(k_{j+1}) \Delta_{l_{j+2}}^T(|k_j - k_{j+1}|)$  is dominated by the multipoles which saturate the triangle inequality, namely  $l_j = l_{j+1} + l_{j+2}$ . The precise form of the bi-spectrum will depend on our model for the  $k$  dependence of the  $\beta_{k_i}$ , of which we can only guess the form, but the characteristic peak for the flattened triangles should be sufficient to distinguish this effect from the usual contributions to non-gaussianity. Note that in the squeezed limit this leading order contribution vanishes, and a subleading term takes over which is less enhanced. Thus this contribution to the three point function should be clearly distinguishable from the squeezed triangles contribution.



## 5. Higher irrelevant operators, N-point functions

Now let us consider higher order operators and their contributions to various N-point functions. It is possible to make fairly general statements based on the scaling with  $1/a$  as to the magnitude of the terms. For the three point function let us first consider the effects of higher order operators of the form  $\frac{\lambda_n}{M^{2n+4}} ((\nabla\phi)^2)^n$ . Naively since these are increasingly irrelevant operators we might expect them to scale with increasing powers of  $1/a$ . However, expanding to cubic order and defining  $\chi = a\delta\phi$  (this is usual since  $\chi$  is pure oscillatory at sub-Hubble scales with no scale factor dependence) we find

$$\sqrt{-g} \frac{\lambda_n}{M^{4n-4}} ((\nabla\phi)^2)^n \Big|_{3\text{-pt}} \approx \frac{1}{a^2} \frac{\lambda_n}{M^2} \left( \frac{\dot{\phi}}{M^2} \right)^{2n-3} (c_1 \chi'^3 + c_2 \chi' (\vec{\nabla}\chi)^2). \quad (5.1)$$

So all the higher order operators will have the same scaling with  $1/a$  and are just additionally suppressed by  $\dot{\phi}/M^2$ . In particular for DBI-inflation and k-inflation we will still have an interaction scaling as  $1/a^2$  and so similar enhancements to those described. Similar statements can be made for the N point functions where we find

$$\sqrt{-g} \frac{\lambda_n}{M^{4n-4}} ((\nabla\phi)^2)^n \Big|_{N\text{-pt}} \approx \frac{1}{a^{2N-4}} \frac{\lambda_n}{M^{2N-4}} \left( \frac{\dot{\phi}}{M^2} \right)^{2n-N} (c_3 \chi'^N + c_4 \chi'^{N-2} (\vec{\nabla}\chi)^2 + \dots). \quad (5.2)$$

To increase powers of  $1/a$  we must go to genuinely higher derivative interactions (rather than powers of first derivatives) such as  $\frac{1}{M^8} (\nabla_a \nabla_b \phi \nabla^a \nabla^b \phi)^2$ . Note that terms containing  $\square\phi$  can be neglected as they are redundant couplings which can be removed by local field redefinitions. Although the correlation functions are not invariant under these redefinitions, it is precisely because they are local that they will not give rise to any interesting effects at late times. Calculating the three point function we have

$$\frac{1}{M^8} \sqrt{-g} (\nabla_a \nabla_b \phi \nabla^a \nabla^b \phi)^2 \Big|_{3\text{-pt}} \approx \frac{1}{M^8 a^5} \ddot{\phi} (\partial_a \partial_b \chi)^2 \chi'' \approx \frac{1}{M^8 a^5} \ddot{\phi} (\partial_a \partial_b \chi)^2 \partial_i \partial^i \chi. \quad (5.3)$$

So we see that this term scales as  $1/a^5$ . Similarly, there are yet higher derivative interactions that we can construct that scale as higher powers of  $1/a$ . However, these naive scalings may not directly represent the enhancements that arise due to the cancellation of numerators evaluated on the flattened triangles. Further more these higher order derivative terms are necessarily additionally  $H/M$  suppressed which typically cancels out the extra  $1/a$  enhancement. Nevertheless the general form of the enhancement effects for the three point function implied by effective field theory will take the schematic form (to first order in  $\beta_k$ )

$$\frac{\Delta \langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle}{\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle} = \sum_i |\beta_{k_i}| f \left( \frac{k_t}{a(\eta_0)M}, \frac{k_i}{k_t}, \frac{k_{i+1}}{k_t} \right) + \mathcal{O}(\beta_k^2), \quad (5.4)$$

where the function  $f$  is a dimensionless function which has a Taylor expansion in its first argument and similar relations holding for the N-pt functions. Thus we see that if  $\beta_k \neq 0$ , the validity of effective field theory requires that  $k/(Ma(\eta_0)) \ll 1$  (for the Taylor expansion to be valid) as we stressed from the outset. With sufficient e-folds this can be violated and so we discover directly the transplankian problem [26], namely the statement that effective field theory does break down at the beginning of inflation if the initial vacuum state is chosen to be nontrivial. However if  $\beta_k = 0$  there is no reason to doubt effective field theory even when  $k > Ma$ . The physical reason is clear; the BD vacuum contains no particles, so that there is nothing that physically carries transplankian energies, and so no transplankian effects should be expected.

## 6. Conclusions

Our main conclusions can be summarized as follows: Choosing an excited gaussian (and Hadamard) initial state at the beginning of inflation can give rise to additional contributions to the bi-spectrum from sub-Hubble interactions. They are significant for flattened triangles  $k_1 = k_2 + k_3$  (and permutations), which in  $l$  space saturate the triangle inequality  $l_1 = l_2 + l_3$  (and permutations). Schematically the corrections from a given operator calculated in an excited state, relative to the equivalent BD state contribution take the form

$$\left. \frac{\Delta\langle\zeta^3\rangle}{\langle\zeta^3\rangle} \right|_{\text{flattened}} \sim |\beta_k| |k\eta_0|^{s+1}, \quad (6.1)$$

where the integer  $s$  depends on the precise scaling of the interaction with inverse scale factor. The specific contribution to  $f_{NL}$  for higher derivative interactions is of the order

$$\left. \frac{\Delta\langle\zeta^3\rangle}{\langle\zeta^3\rangle} \right|_{\text{flattened}} \sim |\beta_k| \left( \frac{k}{a(\eta_0)H} \right)^2. \quad (6.2)$$

The requirement that the excited inflaton quanta do not spoil the slow roll evolution imposes the bound  $|\beta_k| < \sqrt{\epsilon\eta'} \frac{HM_{\text{pl}}}{M^2}$  while the validity of effective field theory bounds  $\frac{k}{a(\eta)} < M$ . Thus the largest expected enhancement factor is

$$\left. \frac{\Delta\langle\zeta^3\rangle}{\langle\zeta^3\rangle} \right|_{\text{flattened}} \sim \sqrt{\epsilon\eta'} \frac{HM_{\text{pl}}}{M^2} \left( \frac{M}{H} \right)^2 = \sqrt{\epsilon\eta'} \frac{M_{\text{pl}}}{H} \quad (6.3)$$

which can easily be of order  $10^4$ . Whilst it may be possible to observe this enhancement in a genuinely three dimensional probe, the CMB bi-spectrum is only sensitive to a two

dimensional projection of this. On converting to  $l$  space we must average over angles which gives a reduction of one factor of  $|k\eta_0|$  so that

$$\left. \frac{\Delta f_{NL}}{f_{NL}} \right|_{\text{flattened}} \sim |\beta_k| |k\eta_0|^s. \quad (6.4)$$

In particular for the higher derivative interactions, the largest expected value is

$$\Delta f_{NL}|_{\text{flattened}} \sim \epsilon \sqrt{\epsilon\eta'} \frac{H M_{\text{pl}}}{M^2} \frac{M}{H} \frac{\dot{\phi}^2}{M^4} \sim 100 \sqrt{|\beta_k|} \frac{\dot{\phi}^2}{M^4} \sim 10^4 |\beta_k|^{5/2}. \quad (6.5)$$

In this particular case we need  $|\beta_k| > 2.5 \times 10^{-2}$  to obtain  $f_{NL} \geq 1$ . Whether this departure from BD can already be detected in the power spectrum, depends on the precise  $k$  dependence of  $|\beta_k|$ . Note that in the extreme case in which  $\beta_k$  is constant over the range of scales probed by the CMB, the only affect of  $\beta_k$  in the power spectrum would be to renormalize its amplitude. On the other hand a stronger  $k$ -dependence could be probed in any spectral tilt present. Nevertheless to get an enhancement relative to the usual result we only need  $|\beta_k| > 10^{-4}$  and so clearly models which already give a large non-gaussian contribution will be the most interesting in terms of seeing the enhancement.

It is clear from our discussion that typically these effects will only be seen if the number of e-folds of inflation is low, i.e. little more than the require  $\sim 55 - 60$ , since we require that the modes corresponding to the observable scales today had physical momenta less than  $M \leq M_{\text{pl}}$  at the onset of inflation. Thus the effect is in the exact opposite region to the transplanckian problem which typically requires  $\sim 70$  e-folds for the physical momenta to reach  $M_{\text{pl}}$ . We can imagine a two-stage inflation scenario however where inflation occurs before this period for a possibly large number of e-folds with the usual Bunch-Davies vacuum, and then some non-adiabatic process excites the vacuum state up to scales set by the Hubble constant at that time  $\sim M$  (e.g. a sudden change in the trajectory of the inflaton in a multidimensional landscape). Then if a second period of lower scale inflation occurred during which the fluctuations relevant for observable scales today were created, these would arise from this excited state and show the signatures we have described.

The essential feature of our result was anticipated independently in Ref. [17] where a general calculation was performed for theories of the form  $p(X, \phi)$ . However these authors did not give any estimate as to the possible magnitude of the effect which we have here emphasized is strongly constrained by backreaction. Furthermore as explained there is little chance of observing the full enhancement for the flattened triangles, but rather only smoothed one which is further suppressed by  $H/M$ . When dealing with theories of the form  $p(X, \phi)$  there is an implicit assumption that one can

take seriously the effective field theory beyond the normal expected range of validity, by taking seriously the effect of higher derivatives. By contrast in this article our perspective has been that even in standard slow roll minimally coupled inflation, these corrections will exist due to standard effective field theory arguments, and we have remained strictly in a regime in which they are small so that standard effective field theory is valid. To properly analyze the case considered in Ref. [17] it will be necessary to redo the backreaction analysis [37].

Irrelevant operators may have very relevant physical consequences, and indeed as we see here, because the relevant contribution to the three point function is naturally suppressed, the first irrelevant higher derivative operator can give a larger correction. For these effects to be observable the cutoff scale of the theory must be less than the Planck scale. Nevertheless, an effect can easily be present for reasonable scales  $H \sim 10^{13-14}$  GeV and  $M \sim 10^{15-16}$  GeV. We may note that super-horizon approaches to the calculation of non-gaussianities [36] will not be able to describe the effects we have considered precisely because they arise from sub-Hubble scale interactions.

From a theoretical point of view, we can use this calculation to get a better understanding of the special nature of the BD vacuum as the true vacuum of interacting quantum fields in de Sitter. What we have shown here can be reinterpreted as the statement that the BD vacuum is self-consistent in that if we choose it to define a Gaussian initial state, although a non-zero three-point function will be generated, it will be highly suppressed both by slow-roll parameters as well as by factors of  $H/M$ . Any initial deviations from the BD state will induce potentially large corrections to the three-point function (although it is important to note that these are still under perturbative control). Keeping the three point function small enough to be consistent with observations then requires a fine-tuning of the initial state (via the  $\beta_k$ 's).

We believe these results also offer a new window to address the transplankian problem, since for the reasons we have explained, higher N-pt functions can be a better probe of the transplankian era than the two point function, provided at least that we start out in an initially excited state. On the other hand it is likely that if the initial state is taken to be the equivalent of the BD vacuum in the transplankian region, these effects will be suppressed at least by  $H/M$ .

In conclusion then we see that the CMB bi-spectrum will be a most useful tool for probing the inflaton initial state, and can contain richer information than the power spectrum alone. Initial state effects will be present in flattened triangles, and a direct observation for an enhancement in these triangles would provide the strongest window into the physics at the beginning of inflation.

## Acknowledgments

We would like to thank Latham Boyle, Cliff Burgess, Claudia de Rham, Olivier Dore, Stefan Hofmann, Justin Khoury and Mark Wyman for useful discussions. R. H. would like to thank the Perimeter Institute for hospitality during some of this work. He also acknowledges support from the Department of Energy through DOE grant No. DE-FG03-91-ER40682. The work of A. J. T. at the Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI.

## References

- [1] D. N. Spergel *et al.* [WMAP Collaboration], *Astrophys. J. Suppl.* **170**, 377 (2007) [arXiv:astro-ph/0603449].  
G. Hinshaw *et al.* [WMAP Collaboration], *Astrophys. J. Suppl.* **170**, 288 (2007) [arXiv:astro-ph/0603451].  
L. Page *et al.* [WMAP Collaboration], *Astrophys. J. Suppl.* **170**, 335 (2007) [arXiv:astro-ph/0603450].
- [2] M. Tegmark *et al.* [SDSS Collaboration], *Phys. Rev. D* **69**, 103501 (2004) [arXiv:astro-ph/0310723].
- [3] G. R. Dvali and S. H. H. Tye, *Phys. Lett. B* **450**, 72 (1999) [arXiv:hep-ph/9812483].
- [4] Peter Svrcek and Edward Witten *JHEP* **0606**, 051 (2006) [arXiv:hep-th/0605206].  
Peter Svrcek “*Cosmological Constant and Axions in String Theory*” [arXiv:hep-th/0607086] (2006).  
M. E. Olsson, *JCAP* **0704**, 019 (2007) [arXiv:hep-th/0702109].  
J. R. Bond, L. Kofman, S. Prokushkin and P. M. Vaudrevange, *Phys. Rev. D* **75**, 123511 (2007) [arXiv:hep-th/0612197].  
A. D. Linde, *In \*Santa Monica 1990, Proceedings, Trends in astroparticle physics\* 105-122*  
R. Holman and J. A. Hutasoit, arXiv:hep-th/0603246.
- [5] N. Kaloper, M. Kleban, A. E. Lawrence and S. Shenker, *Phys. Rev. D* **66**, 123510 (2002) [arXiv:hep-th/0201158].
- [6] P. Creminelli, *JCAP* **0310**, 003 (2003) [arXiv:astro-ph/0306122].
- [7] J. M. Maldacena, *JHEP* **0305**, 013 (2003) [arXiv:astro-ph/0210603].

- [8] V. Acquaviva, N. Bartolo, S. Matarrese and A. Riotto, Nucl. Phys. B **667**, 119 (2003) [arXiv:astro-ph/0209156].
- [9] T. S. Bunch and P. C. W. Davies, Proc. Roy. Soc. Lond. A **360**, 117 (1978).  
N. D. Birrell and P. C. W. Davies, *Cambridge, Uk: Univ. Pr. ( 1982) 340p*
- [10] N. Kaloper, M. Kleban, A. Lawrence, S. Shenker and L. Susskind, JHEP **0211**, 037 (2002) [arXiv:hep-th/0209231].  
B. Greene, K. Schalm, J. P. van der Schaar and G. Shiu, *In the Proceedings of 22nd Texas Symposium on Relativistic Astrophysics at Stanford University, Stanford, California, 13-17 Dec 2004, pp 0001* [arXiv:astro-ph/0503458].  
R. Easther, W. H. Kinney and H. Peiris, JCAP **0508**, 001 (2005) [arXiv:astro-ph/0505426].  
R. Brunetti, K. Fredenhagen and S. Hollands, JHEP **0505**, 063 (2005) [arXiv:hep-th/0503022].  
K. Goldstein and D. A. Lowe, Nucl. Phys. B **669**, 325 (2003) [arXiv:hep-th/0302050].  
H. Collins and R. Holman, Phys. Rev. D **70**, 084019 (2004) [arXiv:hep-th/0312143].  
H. Collins, R. Holman and M. R. Martin, Phys. Rev. D **68**, 124012 (2003) [arXiv:hep-th/0306028].  
C. Armendariz-Picon, JCAP **0702**, 031 (2007) [arXiv:astro-ph/0612288].
- [11] S. Weinberg, Phys. Rev. D **72**, 043514 (2005) [arXiv:hep-th/0506236].  
S. Weinberg, Phys. Rev. D **74**, 023508 (2006) [arXiv:hep-th/0605244].
- [12] M. S. Sloth, Nucl. Phys. B **748**, 149 (2006) [arXiv:astro-ph/0604488].  
M. S. Sloth, arXiv:hep-th/0612138.
- [13] D. Boyanovsky, H. J. de Vega and N. G. Sanchez, Phys. Rev. D **72**, 103006 (2005) [arXiv:astro-ph/0507596].  
D. Boyanovsky, H. J. de Vega and N. G. Sanchez, Nucl. Phys. B **747**, 25 (2006) [arXiv:astro-ph/0503669].  
D. Boyanovsky and H. J. de Vega, Phys. Rev. D **70**, 063508 (2004) [arXiv:astro-ph/0406287].  
D. Boyanovsky, H. J. de Vega and N. G. Sanchez, Phys. Rev. D **71**, 023509 (2005) [arXiv:astro-ph/0409406].
- [14] For an excellent review on non-gaussianity from inflation see: N. Bartolo, E. Komatsu, S. Matarrese and A. Riotto, Phys. Rept. **402**, 103 (2004) [arXiv:astro-ph/0406398].

- N. Arkani-Hamed, P. Creminelli, S. Mukohyama and M. Zaldarriaga, JCAP **0404**, 001 (2004) [arXiv:hep-th/0312100].
- L. Senatore, Phys. Rev. D **71**, 043512 (2005) [arXiv:astro-ph/0406187].
- S. Yokoyama, T. Suyama and T. Tanaka, arXiv:0705.3178 [astro-ph].
- N. Barnaby and J. M. Cline, arXiv:0704.3426 [hep-th].
- D. Battefeld and T. Battefeld, JCAP **0705**, 012 (2007) [arXiv:hep-th/0703012].
- I. G. Moss and C. Xiong, JCAP **0704**, 007 (2007) [arXiv:astro-ph/0701302].
- X. Chen, R. Easther and E. A. Lim, JCAP **0706**, 023 (2007) [arXiv:astro-ph/0611645].
- S. A. Kim and A. R. Liddle, Phys. Rev. D **74**, 063522 (2006) [arXiv:astro-ph/0608186].
- F. Bernardeau, T. Brunier and J. P. Uzan, AIP Conf. Proc. **861**, 821 (2006) [arXiv:astro-ph/0604200].
- F. Vernizzi and D. Wands, JCAP **0605**, 019 (2006) [arXiv:astro-ph/0603799].
- D. H. Lyth and Y. Rodriguez, Phys. Rev. Lett. **95**, 121302 (2005) [arXiv:astro-ph/0504045].
- K. Enqvist and A. Vaihkonen, JCAP **0409**, 006 (2004) [arXiv:hep-ph/0405103].
- C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan and L. Senatore, arXiv:0709.0293 [hep-th].
- C. Cheung, A. L. Fitzpatrick, J. Kaplan and L. Senatore, arXiv:0709.0295 [hep-th].
- J. R. Fergusson and E. P. S. Shellard, Phys. Rev. D **76**, 083523 (2007) [arXiv:astro-ph/0612713].
- [15] D. H. Lyth and D. Wands, Phys. Lett. B **524**, 5 (2002) [arXiv:hep-ph/0110002].
- [16] H. Assadullahi, J. Valiviita and D. Wands, arXiv:0708.0223 [hep-ph].
- J. Valiviita, M. Sasaki and D. Wands, arXiv:astro-ph/0610001.
- M. Sasaki, J. Valiviita and D. Wands, Phys. Rev. D **74**, 103003 (2006) [arXiv:astro-ph/0607627].
- K. A. Malik and D. H. Lyth, JCAP **0609**, 008 (2006) [arXiv:astro-ph/0604387].
- K. Enqvist and S. Nurmi, JCAP **0510**, 013 (2005) [arXiv:astro-ph/0508573].
- C. Gordon and K. A. Malik, Phys. Rev. D **69**, 063508 (2004) [arXiv:astro-ph/0311102].
- N. Bartolo, S. Matarrese and A. Riotto, Phys. Rev. D **69**, 043503 (2004) [arXiv:hep-ph/0309033].
- [17] X. Chen, M. x. Huang, S. Kachru and G. Shiu, JCAP **0701**, 002 (2007) [arXiv:hep-th/0605045].

- [18] A. Gangui, J. Martin and M. Sakellariadou, arXiv:astro-ph/0206023.  
A. Gangui, J. Martin and M. Sakellariadou, Phys. Rev. D **66**, 083502 (2002) [arXiv:astro-ph/0205202].  
J. Martin, A. Riazuelo and M. Sakellariadou, Phys. Rev. D **61**, 083518 (2000) [arXiv:astro-ph/9904167].
- [19] M. Porrati, arXiv:hep-th/0409210.
- [20] M. Porrati, Phys. Lett. B **596**, 306 (2004) [arXiv:hep-th/0402038].
- [21] For recent reviews on effective field theory see for example: C. P. Burgess, arXiv:hep-th/0701053,  
Ira Z. Rothstein, arXiv:hep-ph/0308266.
- [22] E. Witten, arXiv:hep-th/0106109.  
E. Witten, arXiv:0706.3359.
- [23] The “in-in”, closed-time path or Schwinger-Keldysh formalism was first formulated in: J. S. Schwinger, J. Math. Phys. **2**, 407 (1961), and was subsequently developed in a non-equilibrium statistical mechanics context in L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)].  
K. T. Mahanthappa Phys. Rev. **126**, 329 (1962)  
P. M. Bakshi and K. T. Mahanthappa J. Math. Phys. **4**, 12 (1963)
- [24] H. Collins and R. Holman, arXiv:hep-th/0609002.  
H. Collins and R. Holman, Phys. Rev. D **74**, 045009 (2006) [arXiv:hep-th/0605107].  
H. Collins and R. Holman, arXiv:hep-th/0507081.  
H. Collins and R. Holman, Phys. Rev. D **71**, 085009 (2005) [arXiv:hep-th/0501158].
- [25] S. A. Fulling, London Math. Soc. Student Texts **17**, 1 (1989).
- [26] U. H. Danielsson, arXiv:astro-ph/0606474.  
U. H. Danielsson, Phys. Rev. D **71**, 023516 (2005) [arXiv:hep-th/0411172].  
R. H. Brandenberger, arXiv:hep-th/0210186.  
R. Easther, B. R. Greene, W. H. Kinney and G. Shiu, Phys. Rev. D **66**, 023518 (2002) [arXiv:hep-th/0204129].  
U. H. Danielsson, Phys. Rev. D **66**, 023511 (2002) [arXiv:hep-th/0203198].  
A. A. Starobinsky, Pisma Zh. Eksp. Teor. Fiz. **73**, 415 (2001) [JETP Lett. **73**, 371 (2001)] [arXiv:astro-ph/0104043].



- T. Tanaka, arXiv:astro-ph/0012431.
- J. Martin and R. H. Brandenberger, arXiv:astro-ph/0012031.
- J. Martin and R. H. Brandenberger, Phys. Rev. D **63**, 123501 (2001) [arXiv:hep-th/0005209].
- G. Shiu and I. Wasserman, Phys. Lett. B **536**, 1 (2002) [arXiv:hep-th/0203113].
- [27] D. Boyanovsky, H. J. de Vega and N. G. Sanchez, Phys. Rev. D **74**, 123006 (2006) [arXiv:astro-ph/0607508].
- [28] B. R. Greene, K. Schalm, G. Shiu and J. P. van der Schaar, JCAP **0502**, 001 (2005) [arXiv:hep-th/0411217].
- [29] R. Easther, W. H. Kinney and H. Peiris, JCAP **0505**, 009 (2005) [arXiv:astro-ph/0412613].
- [30] P. Creminelli, L. Senatore, M. Zaldarriaga and M. Tegmark, JCAP **0703**, 005 (2007) [arXiv:astro-ph/0610600].
- P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark and M. Zaldarriaga, JCAP **0605**, 004 (2006) [arXiv:astro-ph/0509029].
- [31] B. S. DeWitt, “The global approach to quantum field theory. Vol. 1, 2,” Int. Ser. Monogr. Phys. **114**, 1 (2003). See Chapter 31, Vol. 2 for a formal discussion of the in-in/closed time path/Schwinger-Keldysh formalism, .
- [32] M. Musso, arXiv:hep-th/0611258.
- [33] C. Armendariz-Picon, T. Damour and V. F. Mukhanov, Phys. Lett. B **458**, 209 (1999) [arXiv:hep-th/9904075].
- J. Garriga and V. F. Mukhanov, Phys. Lett. B **458**, 219 (1999) [arXiv:hep-th/9904176].
- [34] E. Silverstein and D. Tong, Phys. Rev. D **70**, 103505 (2004) [arXiv:hep-th/0310221].
- M. Alishahiha, E. Silverstein and D. Tong, Phys. Rev. D **70**, 123505 (2004) [arXiv:hep-th/0404084].
- [35] D. Babich, P. Creminelli and M. Zaldarriaga, JCAP **0408**, 009 (2004) [arXiv:astro-ph/0405356].
- [36] D. S. Salopek and J. R. Bond, Phys. Rev. D **42**, 3936 (1990).
- G. I. Rigopoulos and E. P. S. Shellard, J. Phys. Conf. Ser. **8**, 145 (2005).
- G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, Phys. Rev. D **73**, 083521 (2006) [arXiv:astro-ph/0504508].

G. I. Rigopoulos, E. P. S. Shellard and B. J. W. van Tent, Phys. Rev. D **72**, 083507 (2005) [arXiv:astro-ph/0410486].

G. I. Rigopoulos and E. P. S. Shellard, JCAP **0510**, 006 (2005) [arXiv:astro-ph/0405185].

G. I. Rigopoulos and E. P. S. Shellard, Phys. Rev. D **68**, 123518 (2003) [arXiv:astro-ph/0306620].

C. T. Byrnes, K. Koyama, M. Sasaki and D. Wands, arXiv:0705.4096 [hep-th].

[37] R. Holman and A. J. Tolley, work in progress.